

Capitolo 6

ELABORAZIONE DI IMMAGINI A COLORI

Il colore viene utilizzato nelle immagini digitali per due motivi principali:

- è un descrittore che semplifica l'identificazione di un oggetto e la sua estrazione da una scena
- gli uomini riescono a distinguere migliaia di gradazioni di colore e di intensità, in confronto a solo due dozzine di tonalità di grigio

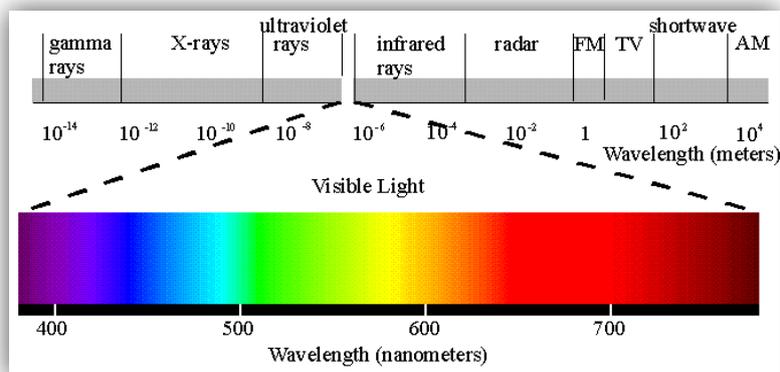
Le due principali *aree di interesse* sono:

- elaborazione *full color*: quando le immagini vengono acquisite con un sensore full-color, come uno scanner o una macchina fotografica digitale
- elaborazione a *falsi colori*: quando si assegna un colore a particolari valori di intensità monocromatici o appartenenti ad una gamma di intensità. Questo tipo di elaborazione viene fatto per semplificare l'analisi delle immagini

Alcune delle tecniche utilizzate per l'elaborazione in toni di grigio possono essere utilizzate anche nell'elaborazione a colori, mentre altre tecniche devono essere riformulate.

Lo **spettro del colore**, in full-color cioè a piena risoluzione cromatica, è formato da sei regioni che non terminano bruscamente, ma ogni colore sfuma gradualmente nel successivo. Questo spettro è creato dal fatto che un prisma di vetro attraversato da un raggio di sole non crea un raggio di luce bianco, ma uno spettro di colori dal viola al rosso.

I colori che gli essere umani percepiscono in un oggetto sono determinati dalla natura della *luce riflessa* dell'oggetto. La luce visibile è composta da una banda relativamente stretta di frequenze nello spettro elettromagnetico [0.43 – 0.79µm].



La **luce acromatica** (senza colore) ha come unico attributo l'*intensità*. La **luce cromatica** invece copre lo spettro elettromagnetico da 400 a 700 nm.

La qualità della luce cromatica viene descritta tramite tre quantità:

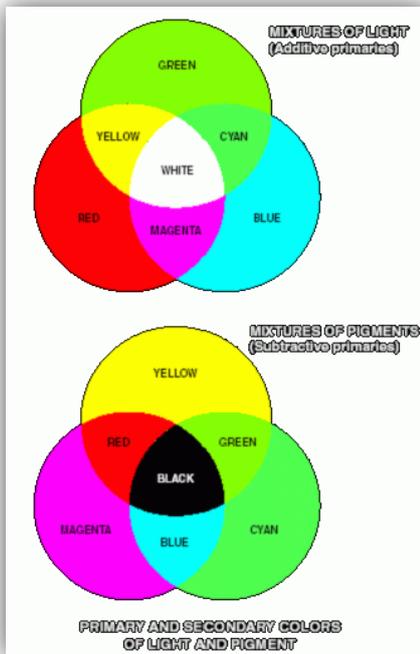
- **radianza**: quantità totale di energia che fuoriesce dalla fonte di luce
- **luminanza**: da una misura della quantità di energia percepita dall'osservatore, misurata in lumen
- **luminosità**: descrittore soggettivo impossibile da misurare

Come abbiamo visto nel capitolo 2 i responsabili della visione del colore sono i coni, che possono essere suddivisi in tre categorie percettive:

- 65% sensibile alla luce *rossa*

- 33% sensibile alla luce *verde*
- 2% sensibile alla luce *blu*

infatti i colori vengono visti come combinazioni variabili dei colori primari RGB. Le tre componenti RGB da sole non possono però generare tutti i colori dello spettro. I colori primari possono essere mescolati per produrre i colori secondari: magenta, giallo e ciano. I colori primari sono usati spesso in modo additivo, quelli secondari in modo sottrattivo.



Le tre caratteristiche utilizzate per distinguere un colore da un altro:

- Luminosità: descrittore soggettivo impossibile da misurare, che ingloba la nozione acromatica di intensità
- Tonalità: attributo associato alla lunghezza d'onda dominante in un insieme di onde luminose, cioè dell'immagine
- Saturazione: purezza relativa della quantità di luce bianca mescolata a una data tonalità. Il grado di saturazione è inversamente proporzionale alla quantità di luce bianca aggiunta. I colori *puri* dello spettro sono pienamente saturi, gli altri sono meno saturi. Tonalità e saturazione prese insieme vengono chiamate *cromaticità*.

I valori **tristimolo** corrispondono alle quantità di *RGB rosso, verde, blu* necessarie per formare un dato colore; queste quantità sono indicate con X, Y, Z . Un colore viene specificato tramite i suoi coefficienti tricromatici definiti:

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

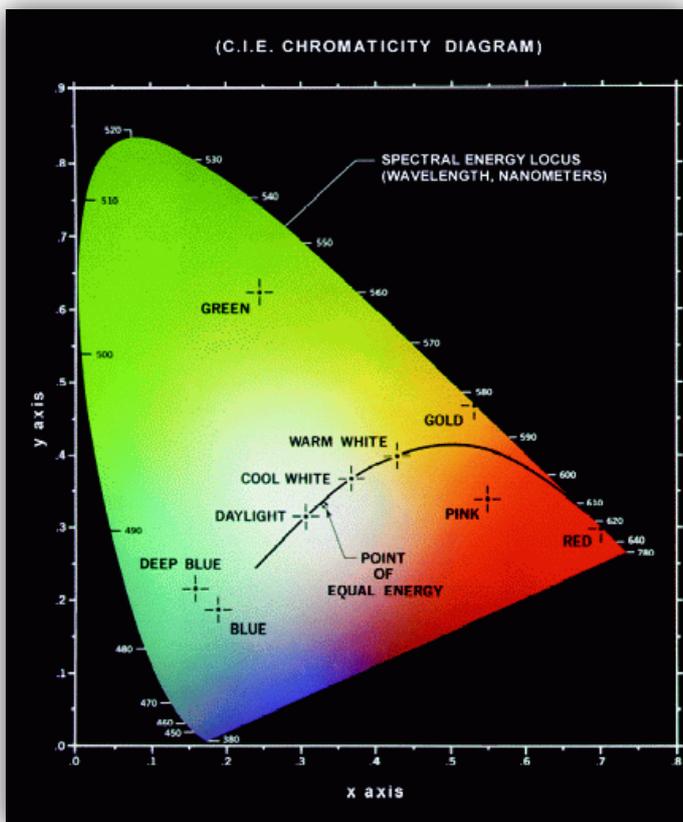
$$z = \frac{Z}{X + Y + Z}$$

quindi $x + y + z = 1$.

Per ogni lunghezza d'onda di luce nello spettro visibile, i valori tristimolo necessari per produrre il colore corrispondente a quella lunghezza d'onda possono essere ottenuti dalle tabelle o curve che sono state compilate a seguito di sperimentazioni.

Il **diagramma di cromaticità** è un altro metodo per specificare un certo colore come funzione di x (rosso) e y (verde). Per ogni valore di x e y , il corrispondente valore di z (blu) si ottiene dall'equazione $z = 1-(x+y)$. In generale se un triangolo interno al diagramma ha i vertici sui tre colori primari fissi RGB (rosso, verde, blu): ogni colore all'interno del triangolo o sul bordo è ottenibile da una combinazione dei tre colori iniziali. Da ciò comprendiamo che con i tre colori primari fissi non si può ottenere tutta la gamma di colori presente nel *diagramma di cromaticità*.

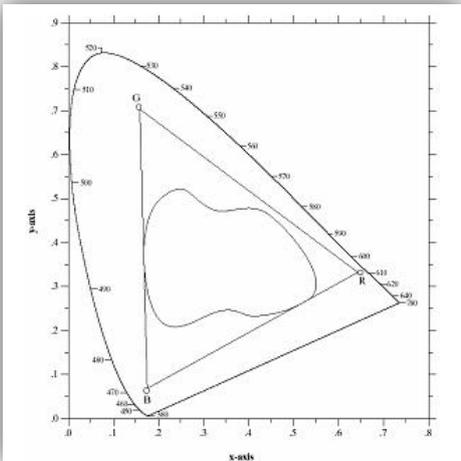
Il diagramma di cromaticità ha la seguente forma:



I colori che si trovano sul bordo del diagramma sono pienamente saturi, man mano che ci si "addentra" nel diagramma aumenta la quantità di luce bianca presente e diminuisce quindi la saturazione del colore. Oltre al triangolo è possibile definire altre figure interne utili, come ad esempio un segmento lineare che unisce due punti qualsiasi nel diagramma definisce tutte le diverse variazioni di colore che possono essere ottenute mediante combinazione lineare dei due colori.

Il **Gamut** (o gamut dei colori) è la gamma di colori racchiusa all'interno di una certa figura geometrica, all'interno del diagramma di cromaticità, che caratterizza un certo dispositivo, ad esempio una stampante a colori o un monitor RGB.

Quindi ad esempio in questa figura:



Il triangolo in dettaglio mostra una tipica gamma di colori prodotta dai monitor RGB. La regione irregolare all'interno del triangolo invece è rappresentativa della serie di colori dei dispositivi di stampa a colori ad alta qualità. Il bordo del range di stampa a colori è irregolare perché la stampa a colori utilizza una combinazione di miscele additive e sottrattive di colori.

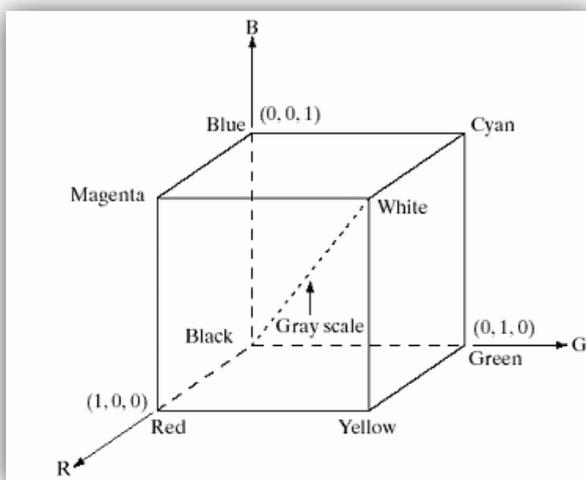
Modelli di Colore

I modelli di colore servono per facilitare e standardizzare la specifica dei colori. E' un sistema di coordinate e di un sottospazio all'interno di quel sistema dove ogni colore viene rappresentato da un singolo punto. La scelta del modello colore dipende dal tipo di applicazione, anche se sono più orientati all'hardware ed alle applicazioni di modellazione del colore:

- *RGB*: monitor a colori
- *CMY(K)*: stampanti a colori
- *HSI*: corrisponde meglio con il modo in cui gli uomini interpretano i colori

Modello RGB

Nel modello RGB ogni colore è rappresentato dalle sue componenti primarie spettrali di *rosso*, *verde* e *blu*, infatti le immagini rappresentate nel modello RGB sono formate da tre immagini, una per ogni colore primario. Si basa su un sistema di coordinate cartesiane, mentre il suo sottospazio è il seguente cubo:



i *valori primari RGB* si trovano su tre spigoli; i *colori secondari* (ciano, magenta e giallo) sono in altri tre spigoli; il *nero* è all'origine e il *bianco* si trova sullo spigolo più distante dall'origine.

I toni di grigio (che hanno uguale valore RGB) vanno dal nero al bianco lungo la linea che unisce i due punti. I diversi colori invece sono punti all'interno o sulla superficie del cubo e vengono definiti in modo vettoriale. La *profondità del pixel* (pixel depth) è il numero di bit utilizzato per rappresentare ogni pixel nello spazio RGB.

I *colori RGB sicuri* sono un sottoinsieme di colori dei quali la riproduzione fedele è indipendente dalle capacità dell'hardware. Il numero minimo di colori che possono essere riprodotti fedelmente da qualsiasi sistema è 256, ma 40 di questi vengono processati in modo diverso dai vari sistemi operativi, mentre i rimanenti 216 sono comuni alla maggior parte dei sistemi, quindi sono diventati gli standard di fatto dei colori sicuri. Vengono utilizzati quando si vuole che i colori visti dalla maggior parte delle persone siano gli stessi. Ognuno di questi 216 colori sicuri è formato da tre valori RGB, ma ogni valore può essere solo 0, 51, 102, 153, 204, 255.

Le facce del cubo che rappresenta i colori sicuri sono le seguenti:

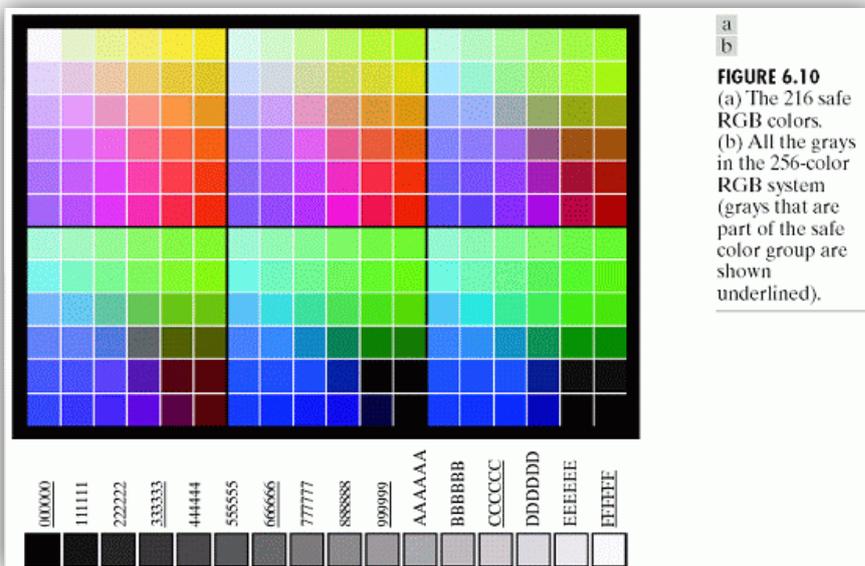


FIGURE 6.10
 (a) The 216 safe RGB colors.
 (b) All the grays in the 256-color RGB system (grays that are part of the safe color group are shown underlined).

Nel cubo corrispondente i colori sicuri sono solo nei piani in superficie (infatti sono le facce). Ogni piano ha 36 colori e l'intera superficie del cubo è coperta da 216 diversi colori.

Concludendo quindi questo modello è utile per la creazione di immagini a colori, ma non per la descrizione del colore.

I modelli CMY e CMYK

Questo modello serve per i dispositivi di stampa a colori che richiedono input di dati CMY o applicano una conversione da RGB a CMY con la seguente operazione:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

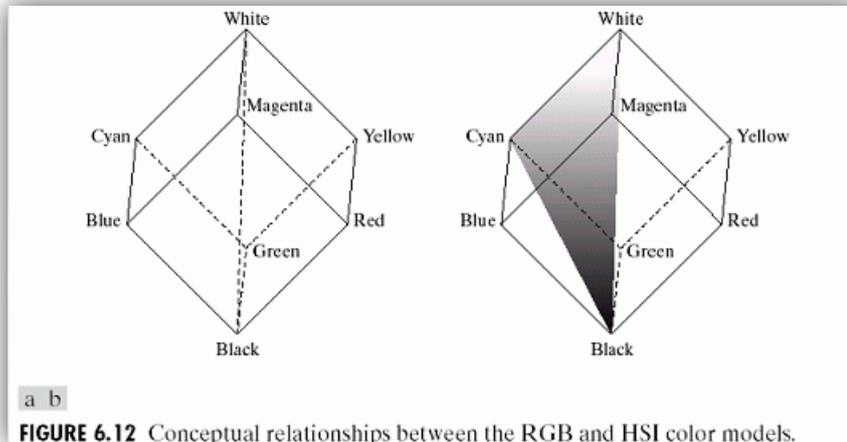
Dove è richiesto che i valori vengano normalizzati nel range $[0, 1]$.

Girando l'espressione è possibile dal modello CMY ottenere i valori RGB.

Il modello *CMYK* (che aggiunge un quarto colore: nero) viene utilizzato per poter stampare un vero nero, che non è possibile ottenere con uguali quantità di pigmenti primari (ciano, magenta, giallo), in quanto producono un colore *simile* al nero.

Il modello HSI

I modelli visti finora non sono molto adatti a descrivere i colori in termini pratici per l'interpretazione umana, per questo motivo è stato introdotto il modello HSI che descrive i colori in termini di tonalità, saturazione ed intensità, quindi si basa su descrizioni dei colori più naturali e intuitive. Il modello divide la componente intensità dalle informazioni relative al colore (tonalità, saturazione) di un'immagine a colori. Questi due cubi permettono di confrontare il modello RGB con quello HSI:



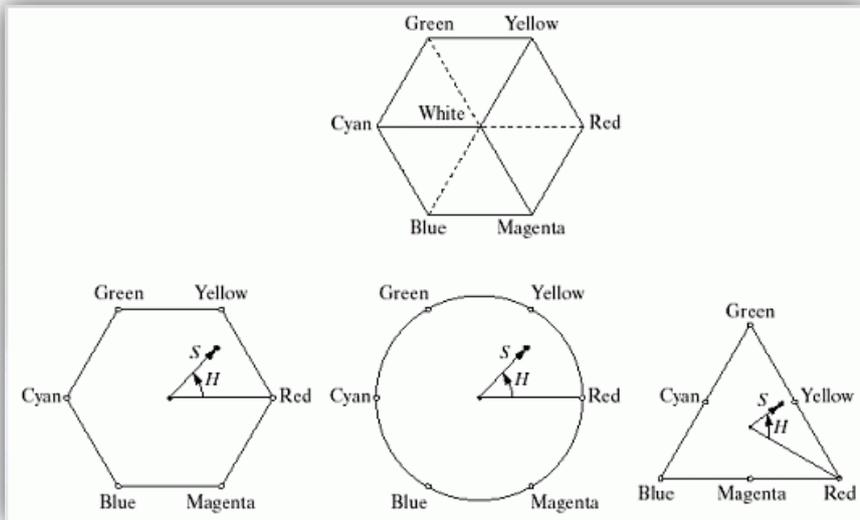
Notiamo che:

- l'intensità si trova lungo la linea che unisce i due vertici bianco e nero, che è verticale ed aumenta lungo questo asse verticale.
Per poter determinare la componente di intensità di ogni punto colore, dobbiamo far passare un *piano perpendicolare* all'asse di intensità che contenga il punto. L'intersezione del piano con l'asse di intensità da un punto con valore di intensità nella gamma [0, 1].
La componente di intensità è data da:
$$I = \frac{1}{3}(R + G + B)$$
- la saturazione di un colore aumenta in funzione della distanza dall'asse di intensità; nello specifico è la lunghezza del vettore dall'origine a quel punto. L'origine è definita dall'intersezione del piano colore con l'asse di intensità verticale.
- la tonalità del punto è determinata dall'angolo rispetto a qualche punto di riferimento. Solitamente un angolo di 0° dall'asse del rosso indica la tonalità 0 e la tonalità aumenta in senso orario a partire da quel punto.
- possiamo convertire qualunque punto RGB in un punto corrispondente nel modello HSI applicando le formule geometriche.
- il concetto più importante è che lo spazio HSI è rappresentato da un asse di intensità verticale e dal luogo dei punti colore che giacciono su piani perpendicolari a questo asse. Dato che i piani si muovono lungo l'asse di intensità su e giù, i bordi definiti da una intersezione di ogni piano con le facce del cubo hanno una forma triangolare o esagonale.
- la forma scelta per il piano HSI può essere un esagono, un triangolo, un cerchio, ma non è un problema perché ognuna di queste forme può essere trasformata in una delle altre due attraverso una trasformazione geometrica.

Riassumendo le *componenti principali* dello spazio HSI sono:

- L'asse di intensità verticale
- La lunghezza del vettore verso un punto colore
- L'angolo che il vettore descrive con l'asse del rosso

Vediamo un'immagine di esempio che rappresenta la tonalità e la saturazione nel modello HSI:



Il punto è un punto colore arbitrario.

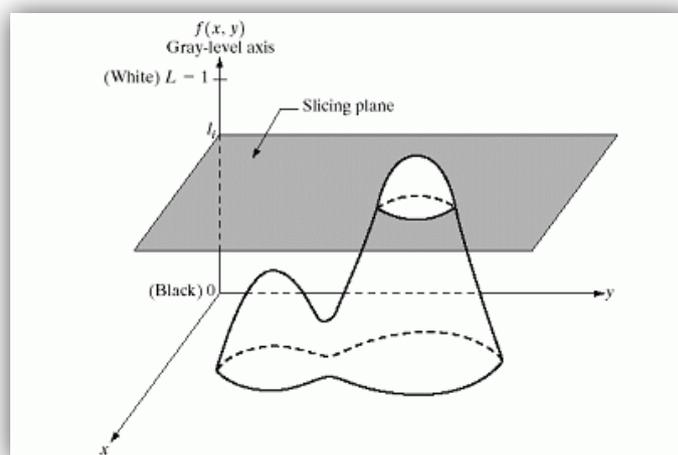
L'angolo dall'asse rosso dà la tonalità e la lunghezza del vettore è la saturazione. L'intensità di tutti i colori in ognuno di questi piani è data dalla posizione del piano sull'asse di intensità verticale.

Elaborazione di immagini a falsi colori

L'elaborazione di immagini a falsi colori è un metodo che consiste nell'assegnare opportunamente l'informazione colore ai valori di grigio. Il termine *pseudo colore* viene utilizzato per differenziare il processo di assegnazione dei colori per immagini monocromatiche da quanto avviene per le immagini *true color*. L'uso principale delle immagini a falsi colori si ha nel settore della visualizzazione e dell'interpretazione visiva da parte dell'uomo di fenomeni a toni di grigio descritti da una o più immagini in sequenza.

Ripartizione delle intensità (slicing)

La ripartizione delle intensità (slicing) è un esempio di elaborazione di immagini a falsi colori. Considerando una funzione 3D come quella in figura:



Questo metodo consiste nel porre dei piani paralleli al piano delle coordinate dell'immagine, quindi ogni piano divide la funzione nell'area di intersezione. Successivamente si assegna un colore ad un lato del piano ed un altro colore all'altro lato, quindi ogni pixel il cui livello di intensità si trova al di sopra del piano sarà

codificato con un colore e ogni pixel al di sotto del piano sarà codificato con un altro colore. Ai pixel che si trovano sul piano stesso può essere assegnato arbitrariamente uno dei due colori.

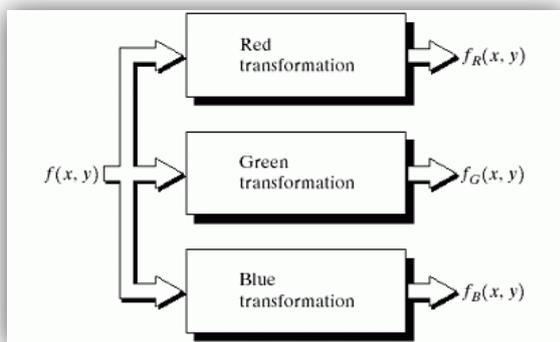
Con questo metodo otteniamo un'immagine binaria il cui aspetto è controllato dalla posizione di I_i sull'asse delle intensità.

Nel caso in cui vengano utilizzati più livelli allora la funzione di trasformazione assume una forma a scala con più gradini.

Questo metodo diventa semplice ed efficace nel caso in cui siano noti i valori di intensità a priori.

Trasformazioni da intensità a colore

Un altro metodo nell'elaborazione di immagini a falsi colori è quello delle trasformazioni da intensità a colore che consiste nell'attuare *tre trasformazioni indipendenti sull'intensità di ogni pixel di input*, quindi ci basiamo sull'intensità non sulla posizione dei pixel. I tre risultati sono visualizzati separatamente nei canali rosso, verde e blu. Alla fine del processo otteniamo un'immagine composta il cui contenuto di colore è regolato dalla natura delle funzioni di trasformazione.



Fondamenti di elaborazione delle immagini full-color

I metodi di elaborazione di immagini full-color possono essere classificati in due categorie:

- Ogni componente viene elaborata individualmente e poi si procede alla combinazione dell'immagine a colori finale
- Si lavora direttamente con i pixel a colori di input.

Dato che le immagini full-color hanno almeno tre componenti i pixel a colori sono dei vettori. Quindi avremo ad esempio:

$$\mathbf{c} = \begin{bmatrix} c_R \\ c_G \\ c_B \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Dove le componenti di \mathbf{c} sono le componenti RGB di un'immagine a colori in quel punto.

Le due categorie non sono sempre equivalenti, non danno sempre gli stessi risultati. Per fare in modo che siano equivalenti bisogna soddisfare due condizioni:

1. Il processo deve essere applicabile sia a vettori che a valori scalari
2. L'operazione su ogni componente di un vettore deve essere indipendente dalle altre componenti

Trasformazioni di colore

Con il termine trasformazioni di colore di si riferisce alle tecniche di elaborazione delle immagini per singola componente e non alle tecniche di conversione degli spazi colore.

Ripartizione di colore (slicing)

Il **Color Slicing**, come nel caso delle immagini in b/n , serve per evidenziare una gamma specifica di colori e risulta utile per separare gli oggetti da ciò che li circonda.

L'idea è quella di visualizzare i colori di interesse in modo che emergano dallo sfondo ed utilizzare la regione definita dai colori come maschera per ulteriori elaborazioni.

Ovviamente le trasformazioni a colori sono più complicate delle loro controparti in scala di grigio. Un metodo semplice per ripartire un'immagine a colori è trasformare i colori al di fuori della gamma di interesse in un colore neutrale non promittente.

Elaborazione di istogrammi

Le **operazioni sugli istogrammi** viste per le immagini in toni di grigio possono essere applicate anche alle immagini a colori.

Ricordiamo che l'equalizzazione di istogrammi, come descritto nel capitolo 3, serve per produrre un'immagine con un istogramma uniforme dei valori di intensità.

Come abbiamo visto prima però le immagini a colori sono formate da varie componenti, quindi bisogna adattare la tecnica utilizzata in scala di grigio a più componenti e/o istogrammi. Per un istogramma comunque è difficile equalizzare le componenti di un'immagine a colori indipendentemente, perché ciò comporterebbe degli errori nel dominio del colore. La soluzione migliore è quella di lavorare nello spazio HSI, nello specifico di distribuire uniformemente le intensità del colore, lasciando invariata la tonalità, cioè i colori stessi. Questo metodo permette di migliorare la qualità di un'immagine andando a modificare soltanto l'intensità e lasciando invariate saturazione e tonalità.

Segmentazione di immagini basata sul colore

La segmentazione, il processo che divide l'immagine in regioni, può essere applicato in due spazi colore:

- Spazio a colori HSI:
si utilizza questo spazio quando si desidera segmentare un'immagine basandosi sul colore e si vuole portare avanti l'elaborazione su singoli piani.
In questo spazio:
 - la *tonalità* rappresenta bene il colore, quindi i suoi valori possono essere utilizzati per segmentare l'immagine
 - la *saturazione* è usata come immagine maschera per isolare ulteriori regioni di interesse rispetto alla tonalità
 - l'*intensità* è utilizzata di meno per la segmentazione in quanto non contiene informazioni riguardanti i colori
- Spazio colori RGB:
spazio nel quale si ottengono i risultati migliori. Dato un insieme di punti campione rappresentativo dei colori di interesse è possibile ottenere una stima del colore medio che si vuole segmentare; quindi l'obiettivo della segmentazione è quello di classificare ogni pixel RGB in una data immagine come un colore appartenente alla gamma specificata o al di fuori di essa; per poter fare questo confronto è necessario avere una *misura di similarità*, in questo caso si utilizza la *distanza euclidea*.

Individuazione di edge a colori

Uno degli strumenti principali che permettono di segmentare un'immagine è l'individuazione dei bordi (edge). Nelle immagini in toni di grigio per trovare i bordi viene utilizzato il gradiente, ma in questo caso non possiamo utilizzarlo perché non è definito per quantità vettoriali, quindi non possiamo calcolare il gradiente sulle componenti RGB singole e poi usare il risultato per formare un'immagine a colori. Al massimo possiamo utilizzarlo lavorando su una singola componente, sempre se ci interessa soltanto individuare i bordi, perché se l'obiettivo è la precisione è necessaria una nuova definizione del gradiente applicabile alle quantità vettoriali.

Siano r, g e b i vettori unitari lungo gli assi R, G e B dello spazio colore RGB e definiamo i vettori:

$$u = \frac{\partial R}{\partial x} \mathbf{r} + \frac{\partial G}{\partial x} \mathbf{g} + \frac{\partial B}{\partial x} \mathbf{b}$$

$$v = \frac{\partial R}{\partial y} \mathbf{r} + \frac{\partial G}{\partial y} \mathbf{g} + \frac{\partial B}{\partial y} \mathbf{b}$$

u e v sono vettori unici ottenuti dalla somma dei tre gradienti.

La direzione della percentuale massima di variazione di $c(x,y)$ (cioè la variazione del gradiente percepita da noi) è data dall'angolo:

$$\theta(x, y) = \frac{1}{2} \tan^{-1} \left[\frac{2g_{xy}}{g_{xx} - g_{yy}} \right]$$

Rumore nelle immagini a colori

Il rumore nelle immagini a colori può essere presente in ogni canale con le stesse caratteristiche, ma è possibile anche che i singoli canali vengano interessati in maniera diversa. La causa principale, di avere diversi livelli di rumore, è la *differenza di illuminazione* disponibile per ogni singolo canale.

Per eliminare il rumore tramite filtraggio di immagini full-color è possibile lavorare sulle singole immagini componenti o direttamente nello spazio vettoriale a colori.

Le caratteristiche principali che emergono dalla presenza di rumore nello spazio RGB e HSI e nel passaggio da uno all'altro sono:

- in un'immagine a colori il *rumore a grana* è meno visibile rispetto che in un'immagine monocromatica
- le componenti di *tonalità e saturazione* sono fortemente degradate, a causa della non linearità delle operazioni di coseno e minimo utilizzate per calcolare le due componenti dallo spazio RGB
- la componente *intensità* risulta più sfocata rispetto alle tre immagini componenti RGB, il motivo è che l'immagine dell'intensità è la media delle immagini RGB
- quando si passa dallo spazio RGB a quello HSI se solo uno dei canali RGB è affetto da *rumore*, la conversione in HSI distribuisce il rumore a tutte le immagini componenti, perché il calcolo delle componenti HSI utilizza tutte le componenti RGB

In questo esempio vediamo le componenti HSI di un'immagine rumorosa:



(a) *tonalità*
 (b) *saturazione*
 (c) *intensità*